

## Electrical connector

**Publication number:** TW305990 (Y)

**Publication date:** 2007-02-01

**Inventor(s):** MA HAO-YUN [TW]

**Applicant(s):** HON HAI PREC IND CO LTD [TW]

**Classification:**

- **international:** H01R13/62; H01R13/62

- **European:** H01R23/68A; H01R13/41; H01R43/20

**Application number:** TW20060205949U 20060410

**Priority number(s):** TW20060205949U 20060410

**Also published as:**

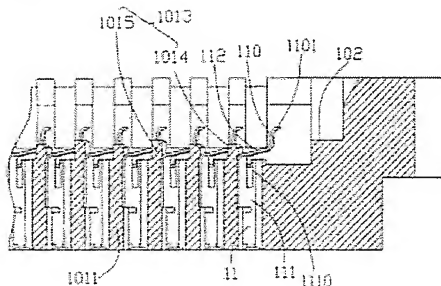
US2007238345 (A1)

US7377792 (B2)

Abstract not available for TW 305990 (Y)

Abstract of corresponding document: US 2007238345 (A1)

An electrical connector comprising an insulative housing defining a mating surface adapted to support an IC package, at least four passageways defined in region within the mating surface and each passageway associated a clapboard arranged between every two adjacent passageways; at least four contacts received in corresponding passageways, respectively, each contact including a contact engaging portion extending upwardly above the mating surface and toward the IC; wherein the clapboard has protrusions extending upwardly therefrom along a direction perpendicular to a bottom surface of the IC package seated on the mating surface, the protrusions comprising first protrusions and second protrusions higher than the first protrusions; the first protrusions being arranged between every two adjacent contacts, allowing the engaging portion of the two adjacent contacts without interference when the IC is seated on the mating surface of the housing.



Data supplied from the esp@cenet database — Worldwide

# 公告本

305990

申請日期	85.8.13
案 號	85109787
類 別	G10 L 5/02

86年3月8日修正補充

305990

(以上各欄由本局填註)

## 發明專利說明書

一、發明 名稱	中 文	具有一聲音元件資料庫之語音合成器
	英 文	"SPEECH SYNTHESIZER HAVING AN ACOUSTIC ELEMENT DATABASE"
二、發明 人	姓 名	1. 伯德 摩比斯 2. 約瑟夫 菲利浦 歐里維 3. 麥克 亞雷罕 譚伯特 4. 珍 皮耶特 文生特
	國 籍	1. 德國 2. 3. 4. 美國
	住、居所	1. 美國新澤西州查坦市希克里廣場25號C6 2. 美國新澤西州華成市維利道101號 3. 美國紐約州紐約市西71街309號1B 4. 美國紐約州布魯克林市魯比路293號
三、申請人	姓 名 (名稱)	美商AT&T公司
	國 籍	美國
	住、居所 (事務所)	美國紐約州紐約市美國大道32號
	代 表 人 姓 名	約翰·J·吉桑

305990

(由本局填寫)

承辦人代碼：
大類：
I P C分類：

A6  
B6

本案已向：

國(地區) 申請專利，申請日期： 案號： ，☐有 ☐無主張優先權  
美 1995.8.16 08/515 887

有關微生物已寄存於： ，寄存日期： ，寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

## 四、中文發明摘要(發明之名稱: 具有一聲音元件資料庫之語音合成器)

一種語音合成方法採用一由在一語音信號期間發生之音標序列建立的聲音元件資料庫。在建立該資料庫時，決定各個包含一對應於一特定音素之音節之音標序列之軌跡。然後根據一對應於不同音素序列之軌跡的集中度確認一誤差容許範圍。資料庫的聲音元件係藉確認對應於沿著緊鄰該誤差容許範圍之相關軌跡的時間點之音標序列內的切斷點而由音標序列之部份形成。依此方式即可能連貫具有一共通連接音素之聲音元件而使在連接音素處之可察覺不連續極小化。本發明亦說明了判定誤差容許範圍的在運算上簡單且快速的方法。

## 英文發明摘要(發明之名稱: "SPEECH SYNTHESIZER HAVING AN ACOUSTIC ELEMENT DATABASE")

A speech synthesis method employs an acoustic element database that is established from phonetic sequences occurring in an interval of a speech signal. In establishing the database, trajectories are determined for each of the phonetic sequences containing a phonetic segment that corresponds to a particular phoneme. A tolerance region is then identified based on a concentration of trajectories that correspond to different phoneme sequences. The acoustic elements for the database are formed from portions of the phonetic sequences by identifying cut points in the phonetic sequences which correspond to time points along the respective trajectories proximate the tolerance region. In this manner, it is possible to concatenate the acoustic elements having a common junction phonemes such that perceptible discontinuities at the junction phonemes are minimized. Computationally simple and fast methods for determining the tolerance region are also disclosed.

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

## 五、發明說明(1)

### 發明領域

本發明概略關於語音合成，更明確地說，本發明係關於一種使用於語音合成中包含聲音元件的資料庫。

### 發明背景

此規則為基礎的語音合成被使用於各種型式的語音合成應用中，包括文字轉語音及聲音回應系統。典型的以規則為基礎的語音合成技術涉及將取自錄音語音之雙音音標序列連貫以形成新字句。此種文字轉語音合成器的一個範例是由本發明讓受者之關係企業製造的TTS系統，該系統描述於由 R.W. Sproat與 J.P. Olive所著 "Text-to-Speech Synthesis"一文中，該文在 AT&T Technical Journal, Vol.74, No.2, pp.35-44(1995年3月/4月)中發表並在此附上供卓參。

一音素(phoneme)相當於語音聲音的最小單位而其作用是分辨各個發音(utterance)。譬如，在英語中，/r/音素相當於"R"字的聲音。音節(phonetic segment)是一音素的特定發音。同樣地，音標序列(phonetic sequence)係相鄰音節的語音間隔。一雙音(diphone)音標序列是一約啓始於一音節的中央部份且約終止於次一音節的中央部份之音標序列。結果，一雙音相當於從一音素至次一音素的轉換。

一般而言，一對應於一音素之音節的中央部份有大致穩定的聲音特性而不會隨時間做大幅變化。所以，形成於二連貫的音標序列之間的接合點處之任何不連續應相當小。但從不同發音取得之連貫音標序列常會產生可察覺的不連續而有害於結果聲音信號之可理解性。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(2)

解決此不連續問題的語音合成方法包括(Academic Press Limited 1995) Computer Speech and Language 一書 1-16 頁由 N. Iwahashi 與 Y. Sagisaka 所著 "Speech Segment Network Approach for an Optimal Synthesis Unit Set"(Iwahashi 等人文獻)及 IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 34, No.2, 264-271 頁 (1986 年 4 月)由 H. Kaeslin 所著 "A Systematic Approach to the Extraction of Diphone Elements from Natural Speech"(Kaeslin 文獻)中, 該二文獻在此附上供卓參。

Iwahashi 文獻的方法使用最佳化技術從預錄語音中選擇可重新合併的雙音音標序列而有較少的不連續或音節間失真。更明確地說, 此方法判斷從錄製的語音中抽取之不同音標序列眾多組合的音節間失真值。其結果的諸失真值接著被利用數學最佳化方法評估以選擇在特定語音內使用的各雙音之整體最佳序列。但此種方法在運算上太過複雜且需特殊的電腦或為人所不愛的長計算時間期間。而且, 雖然雙音音標啓始於一音節的穩態中心並終止於次一音節的穩態中心, 經常會在中心區域內有特殊點在被當作切斷點時會產生造成較差連貫連續性之序列。所以, 音節間失真的降低大幅依賴各音標序列特殊啓始及終止切斷點的選擇品質。這些切斷點一般由人工操作員決定, 該操作員從錄音語音中抽離序列而不知那些切斷點會提供顯著的優點。

Kaeslin 文獻揭櫫一種方法意圖決定最佳啓始及終止切斷點以便降低連貫不連續性。此方法產生所有包含一對應於一特殊音素的音節之雙音音標序列的格式頻率軌跡。格式

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

束

## 五、發明說明(3)

軌跡為構成一發音之測量諧波頻率的隨時間而變的圖示。該方法接著接著根據這些軌跡決定一形心向量。該文獻界定一形心向量為「使在一組軌跡上其本身與最接近點之間的平方和最小...，距離係藉對數區域比率距離測量」的向量。該方法接著從錄音語音中切斷音標序列以在對應於最接近形心向量之軌跡的點之時間點處形成雙音資料庫元件。

但形心向量之決定非常困難且在開始時是靠一人工操作員的"最佳猜測"。基於軌跡的特性，若做了一個不對的"最佳猜測"，則一形心向量會不適當地被決定接近一組區域軌跡，而事實上全部軌跡的真實形心向量卻在別處。使用不適當的形心向量會造成序列切斷點產生完全沒有或無法接受地小的不連續降低度。

所以有一種需要，要有一種聲音音節資料庫建構方法，該方法可自動決定各音節之適當切斷點而大幅降低結果連貫音節之不連續性。

## 發明概述

一種語音合成器採用一種聲音元件資料庫，該種資料庫包括從一語音信號在特殊切斷點抽出之被選定音標序列形成之聲音元件。根據本發明，這些切斷點對應於在一誤差容許範圍內或與其靠近的軌跡時間點。誤差容許範圍應事先選定，使得一最低的期望聲音品質可在一連貫聲音元件內達成，一接合音節之該元件的切斷點對應於該誤差容許範圍的極端部份內之時間點。誤差範圍之定位係根據對應於不同音標序列之軌跡的集中度而定。譬如，誤差容許範

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

束

305990

A7

B7

## 五、發明說明(4)

圖可能為一表達空間的一個範圍，諸軌跡形成於其中，該範圍對應於相當於不同音標序列軌跡之最高集中度範圍。換言之，該範圍與此等軌跡中約最大數目個軌跡交叉或靠近。

所以，本發明依賴藉採用一增強的多樣軌跡來判斷誤差容許範圍之位置而獲致的重大且未預期的益處。此多樣性使本發明能更正確地選擇特殊的音標序列及切斷點以形成聲音元件而達成降低連貫不連續。

根據本發明的一種具體實例，軌跡的表現空白由許多個連續格子覆蓋。在此種具體實例中，有可能採用格子的柵格尋找以藉找出至少一格子的範圍而決定誤差容許範圍，該至少一格子交叉於對應於不同音素序列的大於平均數個軌跡。

根據本發明的另一種具體實例，圍繞沿著一軌跡的各點之範圍內的格子被找出。對每個被找出的格子而言，為該格

(請先閱讀背面之注意事項再填寫本頁)

裝

訂



## 五、發明說明(5)

索引值。然後，解析度值可加或減到經轉換之索引值以判斷格子表列中對應於該特定範圍內諸格子之索引值。接著，表列最長的格子可輕易地被找出以決定誤差容許範圍。

如此，根據本發明可用一在運算上簡單且快速的方法產生一聲音元件資料庫而不需特殊電腦或長的處理時間。此種資料庫的記憶體需求相當小且包含可被連貫成聽來相當自然的合成語音的聲音元件。由於該等聲音元件係使用根據一相關誤差容許範圍決定之切斷點從語音信號中選出，故在連貫期間發生之可察覺不連續數得以減少。

本發明的其他特徵及優點可從下文中的詳細敘述及附圖更清楚地了解。

### 圖式簡述

圖1顯示採用根據本發明之聲音元件資料庫的範例性文字轉語音合成器之示意方塊圖；

圖2A-2C顯示一音節的範例性格式的語音頻譜圖；

圖3顯示根據本發明用來形成圖1之聲音元件資料庫之範例性方法的流程圖；

圖4顯示圖3之方法中使用的音標序列之範例性軌跡圖；  
且

圖5顯示決定圖3之方法中使用的誤差容許範圍之範例性方法的流程圖。

### 詳細敘述

圖1顯示採用根據本發明之聲音元件資料庫5的範例性文字轉語音合成器1。為解釋清楚起見，文字轉語音合成器1

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(6)

的功能組件在圖1中以方塊表示。在這些方塊內執行的功能可藉使用共用或包括-但不侷限於-特定功能積體電路(ASIC)之專用硬體而提供，該硬體包括-但不侷限於-特定用途積體電路(ASIC)或執行軟體的一個處理器或多個處理器。使用"處理器"一詞及其形式不應被視為僅專指能夠執行軟體的硬體，而可為執行對應功能並彼此通訊連絡的各別軟體常式。

在圖1中，資料庫5可能駐在一諸如電腦可讀取記憶體之儲存媒體上，該電腦可讀取記憶體包括類如CD-ROM、軟式磁碟機、硬式磁碟機、唯讀記憶體(ROM)及隨機存取記憶體(RAM)。資料庫5包含有對應於不同的音素序列或包括異音(allophone)之多音的聲音元件。(異音是根據環境語音聲音之音素變體。例如，pit-字的氣音/p/與split一字的非氣音/p/是音素/p/的異音。)

為使資料庫5的大小不致過大，聲音元件一般應對應於有限個音素序列，譬如1到3個音素。聲音元件是音標序列，該等音標序列開始於一音素的約穩態之中心且結束於另一音素的約穩態之中心。有可能將聲音元件以線性預測編碼器(linear predictive coder LPC)參數或數位化語音的形式儲存在資料庫5內，此等形式詳述於例如J.P. Olive所著在1990年ESCA Workshop on Speech Synthesis中提出(25-30頁)的文獻"A New Algorithm for a Concatenative Speech Synthesis System Using an Augmented Acoustic Inventory of Speech Sounds"內，該文獻在此附上供卓參。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

## 五、發明說明(7)

文字轉語音合成器1包括一文字分析器10、聲音元件取回處理器15、元件處理及連貫(element processing and concatenation EPC)處理器20、數位語音合成器25及數位轉類比(D/A)轉換器30。文字分析器10以諸如ASCII格式等可讀取格式接收文字並將文字做語法分析成字且進一步將縮寫與數字轉換成字。然後該等字被根據資料庫5內可用的聲音元件分離成音素序列。這些音素序列接著被傳送至聲音元件取回處理器15。

將字剖析成音素序列及縮寫與數字擴展的方法描述在諸如 K. Church 於 Proceedings of the Second Conference on Applied Natural Language Processing (Morristown, NJ 1988) 之 136-143 頁 的 "A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text"、J. Hirschberg 所著 Artificial Intelligence(1993)第 63 卷 第 305-340 頁 的 "Pitch Accent in Context: Predicting International Prominence From Text"; R. Sproat 所著 Computer Speech and Language(1994)第 8 卷 第 79-94 頁 "English Noun-Phrase Accent Prediction for Text-to-Speech"; C. Coker 等人在 Proceedings of the ESCA Workshop on Speech Synthesis(1990)之 第 83-86 頁 的 "Morphology and Rhyming: Two Powerful Alternatives to Letter-to-Sound Rules for Speech" 中，這些文獻均在此附上供卓參。

文字分析器10進一步判斷各個音素序列的期間長度、幅度及基本頻率並將這些資訊傳送給EPC處理器20。判斷期間長度的方法包括在諸如 J. van Santen 所著 Computer

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(8)

Speech and Language(1994)第8卷第95-128頁之"Assignment of Segmental Duration in Text-to-Speech Synthesis"中所述的方法，該文獻在此附上供卓參。判斷一音素序列之振幅的方法描述於諸如 L. Oliveira在 ESCA EUROSPEECH-93 (1993)的第99-102頁之"Estimation of Source Parameters by Frequency Analysis"中，該文獻亦在此附上供卓參。一音素的基本頻率亦可稱為音調或該音節的音調抑揚。決定基本頻率或音調的方法描述於諸如 M. Anderson等人在 Proceedings of the International Conference on Acoustics, Speech and Signal Processing(聖地牙哥1984)第1卷第2.8.1-2.8.4的"Synthesis by Rule of English Intonation Patterns"中，該文獻也在此附上供卓參。

聲音元件取回處理器15從文字分析器10接收音素序列，然後由資料庫5中選擇並取回對應的適當聲音元件。聲音元件選擇方法敘述於諸如上文所提Olive參考文獻中。被取回的聲音元件接著被聲音元件取回處理器15傳送至EPC處理器20。EPC處理器20修改各個接收到的聲音元件，其修改方法是根據從文字分析器10接收到的對應資訊調整聲音元件之基本頻率及振幅，並插入適當的期間長度。EPC處理器20接著將經修改的聲音元件連貫成對應於文字分析器10的文字輸入之聲音元件串。EPC處理器20的連貫方法描述於上文所提的Oliveira參考文獻中。

由EPC處理器20產生之聲音元件串被提供予數位語音合成器25以產生對應於該聲音元件串之自然語音的數位信號

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(9)

。數位語音合成的範例性方法亦描述於上文所提的Oliveira參考文獻中。由數位語音合成器25產生的數位信號被提供予D/A轉換器30以產生對應的類比信號。此等類比信號可被提供予放大器及揚聲器(未顯示)以產生聽來自然的合成語音。

音標序列隨時間而變的特性曲線可用數種表示法表示，包括話音素(formant)、振幅及否定頻譜表示法，包括逆頻譜表示法或任何由LPC導出之參數。圖2A-2C顯示對應於從音節/p-i/的錄音語音中取出之音素/i/的音節之不同話音素頻率或話音素F1、F2及F3的語音頻譜圖100A、100B及100C。話音素F1-F3是表現人類發話者的發音部位之不同測量諧振頻率的軌跡。不同測量諧振頻率的話音素一般是根據被相關話音素包含之頻譜能量而命名為F1、F2、...。

話音素頻率視發音部位的形狀和大小而定。改變發音部位之形狀可形成不同的聲音。所以，當發音部位形狀在音節/i/的發音期間改變時，語音信號之頻譜特性隨時間而變，如圖2A-C所示般。音素/i/的三個話音素F1、F2及F3的顯示僅為舉例說明用。請注意根據一特定語音音節的發音部位形狀可有不同數目種發音。發音和語音的其他表示法的更詳細敘述請見L.R. Rabiner與R.W. Schafer所著"Digital processing of Speech Signals" (Prentice-Hall, Inc. NJ, 1978)中，該文獻在此附上供卓參。

如上文中有關圖1之說明，儲存在資料庫5內之聲音元件對應於音標序列，該等音標序列開始於一音素的約中心部

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(10)

份且終止於另一音素的約中心部份。在二連貫的聲音元件的接合音素處諸如頻譜成份等特性差異會產生不連續而造成合成的語音易理解或難以了解。但在對應於音素中心範圍的音節範圍內常有特別的切斷點在穩態特性範圍內可被用來產生聲音元件而獲致連貫不連續的降低。圖2A-C內各個軌跡F1-F3代表在特殊音素中央範圍處之音標序列特性。有需要要在音標序列內選擇切斷點以形成可使連貫不連續為最小的聲音元件。

圖3顯示一種根據本發明之範例性方法200，本方法從一語音信號中選擇特別的音標序列並決定被選出之音標序列的對應切斷點以形成資料庫5的聲音元件。根據該方法200，包含一對應於一特別音素之音節的音標序列在步驟210中被從一語音信號的一間隔中找出。每一音標序列應對應於一至少有二音素之序列。語音信號有可能從錄音語音或直接由人類發聲者獲得。此外，若語音信號的來源是錄音語音，則該錄音語音可進一步加以處理以產生一分段且加標籤的語音信號而有助於方法200的運作。一分段且加標籤的語音信號是一語音信號，其對應音標序列被加標籤且諸序列間的約略邊界被找出。

然後在步驟220內對對應於該特定音素之各個音標序列的至少一部份決定軌跡。軌跡是該音標序列之該部份的至少一個聲音特性隨時間變化的一種表示法。軌跡可能是表示該聲音特性的一獨立序列或該聲音特性在一段期間內的連續表示法。可被用做軌跡之適當聲音特性範例包括頻譜表

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(11)

示法，例如像話音素頻率、振幅與頻譜傾斜表示法及LPC表示法。其他不論是以頻率為基礎的或其他的聲音特性可根據本發明被用做軌跡。圖2A-C中各顯示一單獨話音素頻率表示法的範例性軌跡。

在步驟220中，軌跡依表示空間(representational space)法決定。如本處所使用者，一表示空間是一區域，一軌跡可在該區域中被描述為可表現該軌跡特性之參數的函數。譬如，圖2A中所示的單一話音素軌跡的表示空間被描述為頻率的時間函數。有可能根據一特定音標序列的二個或更多個話音素頻率形成一單一軌跡。對此一軌跡而言，表示空間對每一被表示的正式頻率都有一軸。沿著各軌跡的頻率有可能被對應時間標示於此等頻率在音標序列發生的時間。例如，一二話音素頻率軌跡會被形成於二度空間中的曲線，其中曲線點的對應時間以5毫秒間隔標示。

軌跡在表示空間內決定之後，在步驟230中根據對應於不同音素序列之軌跡的集中度決定一誤差容許範圍的位置。誤差容許範圍是一在N度表示空間內之N度空間範圍，其交叉或最靠近於一對應於不同音素序列的相當高集中度的軌跡。譬如，誤差容許範圍可能是一交叉或最靠近於一對應於不同音素序列的最大數目個軌跡的範圍。誤差容許範圍的大小應予事先決定以在連貫聲音元件時達成一最低可用的聲音品質，其中一接合音素之切斷點對應於誤差容許範圍的極端點內之時間點。決定適當誤差容許範圍的特別方法在下文中參照圖4及5詳述。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明(12)

在決定誤差容許範圍位置之後，接著在步驟240中選擇特殊的音標序列以根據對應軌跡與誤差容許範圍接近的程度形成聲音元件。例如，若在語音範圍中有數個音標序列對應於相同的音素序列，則其對應軌跡最接近或在誤差容許範圍內之音標序列被選擇以便形成聲音元件。

當音標序列在步驟240內被選擇之後，接著在步驟250中，相關切斷點被決定於音標序列中以獲得期望的聲音元件。該等切斷點對應於沿著大致最接近誤差容許範圍或在誤差容許範圍內之軌跡的時間點。最後，在步驟260中，根據被選擇的音標序列及其對應切斷點形成聲音元件。若在步驟210內找出的所有音標序列都要形成聲音元件，則不論是否因為僅有一個音標序列存在各期望之音素序列的語音信號內，步驟240均可予省略。

根據本發明，誤差容許範圍的位置係根據對應於不同音素序列之軌跡決定。依此方式，本發明藉使用少於來自語音信號之音標序列的總軌跡數之軌跡決定誤差容許範圍之位置而獲致較高的多樣性。此多樣性使本發明能更正確地選擇特殊音標序列及切斷點以形成聲音元件而達成較低的連貫不連續。若一誤差容許範圍的位置是一對應於不同音素序列之軌跡的最高集中度，則該聲音元件會產生相當高聲音品質的合成語音。但若稍差一點的聲音品質可接受，則根據本發明可使用較最高軌跡集中度為低的誤差容許範圍。

根據方法200決定誤差容許範圍的一種範例性技術是將軌

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉



## 五、發明說明 ( 13 )

跡被決定於其內的表示空間分割成各個相關格子，並且找出具有至少一最低期望軌跡集中度位準的特別格子或範圍。根據本技術之方法200的範例性作業現將參照圖4中所示範例性軌跡圖300加以描述。參考圖3，包含對應於音素/i/之音節的音標序列於步驟210中在錄音語音的一時段中被找出。音標序列對應於音素序列/lid/，/lik/，/mik/，/gim/，/din/且5個音標序列對應於音素序列/kit/。可由這些音標序列形成之聲音元件包括[l-i]，[i-d]，[i-k]，[m-i]，[g-i]，[i-m]，[d-i]，[i-n]，[k-i]及[i-t]。雖然圖4之討論考量的是雙音聲音元件的建構，但請注意更大音素序列的聲音元件可根據本發明藉執行圖3之方法200於對應較大音標序列的特定邊界音素上而建構。

對在步驟210內找出的各個音標序列而言，各個音標序列在步驟220中形成二話音素軌跡。圖4中所示軌跡圖300舉例說明這些軌跡在一二話音素表示空間中被分割成許多個格子310。在圖4中，每個軌跡被標示以其對應音素序列的同質體。譬如，軌跡305被由對應於音素序列/lid/之音標序列決定，故被標示為"LID"。來自被用來產生圖1之資料庫5的語音信號之部份的音素序列/kit/之5種狀況被標示為"KIT1"到"KIT5"以利討論。該圖所示各個二話音素軌跡表示在一特定時點上相關音標序列之話音素F1的頻率值相對於對應話音素F2的頻率值的繪圖。

話音素F1與F2之頻率分別顯示於X與Y軸上。沿著軌跡的時間特定點可被表示為對應標示如軌跡305上所示般。圖4

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

## 五、發明說明 ( 14 )

中二維軌跡表示法僅為方便討論及舉例說明用而非對本發明設下任何限制。有可能使用其他N維表示法，包括例如三話音素或四話音素表示法來表示具有一母音做為該特定音素的音節，及振幅及頻譜傾斜表示法來表示具有一子音做為該特定音素的音節。

僅為了易於顯示及說明，表示空間內格子310的格子大小被設定為誤差容許範圍期望大小的1/4。當誤差容許範圍大小實質上不大於格子大小時，設定格子大小為期望誤差容許範圍大小的倍數就很方便。根據圖3之方法200的步驟230，誤差容許範圍之決定是根據被對應於不同音素序列之軌跡交叉的範圍決定。所以，若一由2x2陣列格子310構成的誤差容許範圍被判定其大小足以產生期望的最低聲音品質，則被最大數目個此種軌跡交叉的範圍320即為誤差容許範圍。

決定具有最大數目個此種軌跡交叉之格小的一種方法例如是執行在表示空間內的格子柵格尋找。根據此種方法，圖4之各格子310被檢查，且與該格子交叉對應於不同音素序列之軌跡數目，或圍繞該格子310之預定格子解析範圍被決定。例如，對軌跡LID與MIK而言，對應於格子330之不同音素序列的軌跡交叉數為2。下文中將參考圖5詳述一種在計算上較簡單且較快速之決定具有對應於不同音標序列的最大數目個此種軌跡交叉之格子的方法。

回頭參考圖3之方法200。在軌跡決定之後，接著在步驟240中，根據靠近誤差容許範圍320之對應軌跡選擇特別的

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

## 五、發明說明 ( 15 )

音標序列以形成聲音元件。若對一特定音素序列僅含入資料庫5內的一個聲音元件以使資料庫所需空間極小又使語音合成器設計能簡化將是很有利的。所以，音標序列/lik/或/lid/二者之一可被選擇來形成聲音元件[l-i]，且音標序列/lik/或/mik/二者之一可被選擇來形成聲音元件[i-k]。此外，音素序列/kit/的5個音標序列之一可被選擇來形成聲音元件[k-i]與[i-t]。但一採用更大資料庫之更複雜語音合成器有可能根據語音合成應用用途而使用多重音標序列於一特定音素序列上。在建構此種資料庫時，可從對應於一特定音素序列之從語音信號析出之一個以上及多達全部的音標序列選擇出來以形成聲音元件。

若一特定音素序列要有一聲音元件存在於資料庫5內，則從對應於形成該聲音元素之相同音素序列的許多個音標序列中找出特別的一個可根據對應軌跡靠近誤差容許範圍的程度決定。譬如，對聲音元件[l-i]而言，會選擇其軌道LID與誤差容許範圍320交叉之"LID"音標序列，而不會選擇其軌道LIK不與誤差容許範圍320交叉之"LIK"音標序列。同樣地，聲音元件[i-k]會選擇音標序列"MIK"而不選擇音標序列"LIK"，其原因大致相同。依相同的方法，對聲音元件[k-i]與[i-t]二者而言，會選擇對應於軌跡KIT5的音標序列而不選擇其他相關的音標序列"KIT"。

同時，由於聲音元件一般可在二邊界性音素處連貫，故選擇用來形成聲音元件的特定音標序列應根據該二邊界性音素之軌跡接近程度。所以，特別的音標序列"MIK"或

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

裝

## 五、發明說明(16)

"LIK"會被選擇以形成聲音元件[i-k]，因為該二音標序列之軌跡在整體上最接近邊界性音素/i/與邊界性音素/k/的誤差容許範圍。

有時候對應於相同音素序列之音標序列的軌跡不會是最靠近其邊界性音素的二相關誤差容許範圍者。此種情況發生於當音標序列之來源為包含該音素序列的二個不同字時。在此情況下，較佳的方法是選擇其軌跡具有整體最佳品質的音標序列。選擇此種音標序列的一種範例性方法是根據一特殊品質衡量值各指定一值予各個音標序列以評量諸音標序列有關對應的邊界性音素。評量為整體最佳的音標序列將會被用來形成該聲音元件。

再回頭參考圖3之方法200，聲音元件之音標序列被選擇之後，在步驟250中選擇被用以形成該聲音元件之音標序列切斷點。例如在圖4中，切斷點之選擇係根據在誤差容許範圍320內之相關軌跡中的時間點決定。對那些與誤差容許範圍320交叉的軌跡而言，被選擇之切斷點應宜為沿著軌跡大約最靠近誤差容許範圍320之中央點340的時間點。譬如，圖4中軌跡305上最靠近中央點340的時間點是160毫秒。所以，聲音元件/i-k/是根據開始於時間160毫秒的對應音標序列。

對於諸如軌跡LIK等不與誤差容許範圍320交叉之軌跡而言，切斷點仍應為沿著軌跡最靠近誤差容許範圍中央點340之時間點。故若選擇音標序列"LIK"來形成聲音元件，則適當的切斷點將對應於軌跡LIK上時間點350。請注意當使用

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

錄

## 五、發明說明 ( 17 )

此音標序列來形成聲音元件時，將會在音素/i/處產生相對較大的不連續。所以，最好是能獲得音素序列/lik/的其他語音段落以決定其是否為形成該聲音元件的較佳待選語音段落。

在圖3的方法200中，在切斷點於步驟250中被決定之後，聲音元件根據被選定的音節和被決定的切斷點形成。聲音元件可留存在圖1之資料庫5中，其形式可類如數位語音信號或對應於開始及結束於相關切斷點之音標序列的LPC參數。而且，較長的序列可連同對應於相關聲音元件之特別切斷點的開始及結束值儲存在資料庫5中。圖1之聲音元件讀回處理器15則會根據這些值從這些較長序列之抽出適當的聲音元件。請特別注意的是用於資料庫5之特別組織性方法不應被視為一種限制，且任何組織均可被用以儲存根據本發明形成之聲音元件。為了要合成一特別語言的多種話音素，應產生該種語言的所有基本音素序列之聲音元件。

根據本發明的新穎的使用高度多樣化軌跡來決定誤差容許範圍的位置結果造成在連貫時產生較小不連續的聲音元件。例如在圖4中，範圍360對應於根據所有軌跡之範圍，且交叉於或最靠近於源自音素序列/kit/的5個軌跡之整體最大數目個此種軌跡。但可見到軌跡LTD與MIK上最靠近範圍360的時間點在連貫相關聲音元件時會產生相當大的不連續。相對地，誤差容許範圍320未被音素序列/kit/的多重狀況扭曲，且所有被選擇之軌跡至誤差容許範圍320間的對應

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

## 五、發明說明 ( 18 )

距離遠較小而將使任何相關不連續極小化。

圖5顯示根據本發明使用於圖3中步驟230內決定對應於不同音標序列之具有最大數目個軌跡交叉的格子之範例性方法400。為了便於討論起見，每個軌跡在圖5中用一獨特的整數稱呼而不像圖4中使用相關音標序列標示法。例如，圖4中所示的9個軌跡在圖5中被稱為軌跡1-9。此種軌跡標示法符合使用於資料結構表示法中諸如陣列中或表列中的慣用指向器。

根據方法400，一整數N與許多個表列LIST<sub>i</sub>在步驟410中被初始化為零。該等許多個表列LIST<sub>i</sub>中表列的數目i對應於表示空間中的格子數。接著整數N在步驟420中被增加數值。然後，對軌跡N內的各時間點，在步驟430中找出在圍繞相關時間點之解析範圍內的格子。為方便起見，解析範圍的大小可相同於誤差容許範圍者。但根據本發明，解析範圍的大小亦可依需要而為不同的大小。若選擇的解析範圍是由一2x2格子陣列涵蓋的區域，則圖4中圍繞軌跡305在時間0.095毫秒處之時間點505的解析範圍將包括被外框線510圍繞的格子511、512、513及514。

在解析範圍內的格子於步驟430內找出之後，被找出的格子之相關表列LIST<sub>i</sub>被用對應軌跡N的音素序列名稱更新。而且，在步驟440中，若該音素序列名稱不是已出現在該格子的表列上，則其僅被加至表列。所以，假設在上述範例中名稱"LID"未出現在格子511-514的表列LIST<sub>i</sub>中，則這些格子的表列LIST<sub>i</sub>，將被用該名稱更新。沿著軌跡

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

## 五、發明說明 ( 19 )

305的其他時間點在解析範圍內的格子之表列LIST<sub>i</sub>亦會依大致類似方法用名稱"LID"更新。

當一特定軌跡N被找出之解析範圍內所有格子均在步驟440中更新之後，本方法在步驟450中判斷整數N是否等於軌跡的總數。若本方法判定N不等於軌跡總數，則方法400根據次一軌跡N的時間點執行步驟420-440以更新表列LIST<sub>i</sub>。若本方法判定N等於軌跡總數，則所有的軌跡均已處理完成，且解析範圍內全部表列LIST<sub>i</sub>均已更新完成，方法400前進到步驟460。在步驟460中，誤差容許範圍由在對應表列或諸表列LIST<sub>i</sub>中具有最大數目個名稱的格子或諸格子的範圍決定。因為方法400僅檢驗並更新在軌跡時間點之解析範圍內的諸格子，所以其在運算上較分別檢驗每個格子的柵格尋找方法便簡單且快速。

在方法400中，步驟430首先檢驗一特定軌跡的時間點在解析範圍內所有的格子，然後在步驟440中更新對應的格子表列。但請注意圖4中所示諸步驟的順序僅為舉例說明用而非對本發明之限制。這些步驟的順序可用各種不同方式施行，包括在一表列LIST<sub>i</sub>的相關格子被判定為在一特定軌跡時間點的解析範圍內之後，立刻更新該表列。

在一替代性具體實例中，具有最長表列LIST<sub>i</sub>之格子的等位體可藉儲存並更新具有最長表列LIST<sub>i</sub>之等位體及對應最大表列長度而在整個格子表列更新程序中被維護。當各格子表列被更新時，包含在該表列中的名稱之總數目超過被儲存格子同位體之名稱數目，則被儲存之格子同位體

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

錄

## 五、發明說明(20)

及最大表列長度可據以被更新。依此方式，對應於誤差容許範圍之格子的同位體可在處理前一軌跡的最後一個時間點時即已知而不需任何進一步的處理步驟。

若諸格子表列被加索引，加索引的形式譬如是具有整數值的資料結構指派格子位置於表示空間內，然後可採用一在運算上簡單且較快速的方法。例如，圖4中格子310之格子表列可依一對應於其X與Y座標之方法加索引。然後，轉換值可被用以將軌跡時間點值轉換成表示根據被加索引之格子該等時間點的相對座標位置的索引值。接著，解析度值被加到經轉換的索引值及由其減去以找出該點解析範圍內諸格子之索引數目。然後，解析範圍內各個格子之表列LIST<sub>i</sub>被據以更新。

所以，對圖4中所示範例而言，圖4中軌跡305之時間點505的話音素F1及F2頻率值可乘以轉換因素以獲得經轉換的值 $x=3.5$ 及 $y=3.5$ ，表示其分別在X與Y方向的第三與第四格子之間。所以，若解析範圍是一 $2 \times 2$ 格子陣列，則 $\pm 1$ 的解析值須被加至經轉換的值並四捨五入到要產生之最靠近位置，以使解析範圍510內諸格子的格子表列具有分別對應於格子511-514的座標(3,3)，(3,4)，(4,3)及(4,4)，且會被用音素序列名稱"LID"更新。

雖然上文中詳細描述本發明的數種具體實例，但可在不背離其教導的情形下做許多修改。所有這些修改均旨在包含於下文中申請專利範圍內。譬如，雖然上文中用二維矩形格子及誤差容許範圍來描述本發明，但是有可能使用任

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉



## 五、發明說明 ( 21 )

何N維封閉形狀於格子及範圍上，而符合包括立方體、盒狀體、球狀體及偏球狀體等的N維表示空間。而且，本發明在包含文字轉語音及聲音回應系統的各種語音合成應用中特別有用。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

錄

## 六、申請專利範圍

1. 一種產生合成語音的方法，該種方法包括一聲音元件資料庫，該資料庫包含用來連貫以產生合成語音的聲音元件，該聲音元件資料庫藉下列諸步驟建立：

對於對應於包含在發生於一語音信號的一期間內之許多個音標序列中的特定音節的至少一個音素而言，

根據對應於和一誤差容許範圍交叉的不同音素序列之音標序列的軌跡集中度，決定一表示空間內一誤差容許範圍之相對位置，其中每一軌跡代表包含該特定音節之相關音標序列的至少一部份之聲音特性；及

根據諸時間點與誤差容許範圍接近的程度，藉著找出沿著對應軌跡上各相關時間點處之音標序列內的切斷點，由音標序列形成聲音元件。

2. 根據申請專利範圍第1項之方法，該方法尚包括根據對應諸軌跡接近誤差容許範圍的程度從許多個音標序列中選擇至少一個音標序列的步驟，該等許多個音標序列有對應於一特定音素序列的部份，其中一聲音元件由被選定之音標序列的部份形成。
3. 根據申請專利範圍第1項之方法，其中該形成聲音元件之步驟沿著對應軌跡的相關時間點處找出每一音標序列約略最靠近誤差容許範圍或在誤差容許範圍內的切斷點。
4. 根據申請專利範圍第3項之方法，其中該形成聲音元件之步驟沿著對應軌跡的相關時間點處找出每一音標序列約略最靠近誤差容許範圍中心點的切斷點。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 六、申請專利範圍

5. 根據申請專利範圍第1項之方法，其中一特別語言之每一預測的音素序列均形成一聲音元件。
6. 根據申請專利範圍第1項之方法，其中諸軌跡係根據諸音標序列的諸話音素形成。
7. 根據申請專利範圍第1項之方法，其中該等軌跡係根據一種三話音素表示法，且該表示空間為一三話音素空間。
8. 根據申請專利範圍第1項之方法，其中該表示空間為一包括許多個連續N維格子的N維空間，且其中該決定誤差容許範圍之步驟尚包括執行一柵格尋找以決定一與對應於不同音素序列之約最大數目個軌跡交叉的至少一個格子之範圍。
9. 根據申請專利範圍第1項之方法，其中該表示空間為一包括許多個連續N維格子的N維空間，且其中該決定誤差容許範圍之步驟包括：
 

找出那些沿著每一軌跡在一圍繞諸時間點之解析範圍內的格子；

對在解析範圍內被找出的每一格子，若一對應於該軌跡之音素序列的識別體未出現於該格子的表列內，則用該識別體更新一為該格子維護的表列；及

決定對應於至少一個具有大於其表列上識別體平均數個識別體之格子的誤差容許範圍。
10. 根據申請專利範圍第9項之方法，其中該找出那些在一解析範圍內之格子的步驟包括處理沿著諸軌跡之諸時間點及更新關聯於在對應解析範圍內的諸格子之表列。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 六、申請專利範圍

11. 根據申請專利範圍第9項之方法，其中該解析範圍與該誤差容許範圍的大小相同。

12. 根據申請專利範圍第1項之方法，其中該表示空間為一包括許多個連續N維格子的N維空間，且其中該決定誤差容許範圍之步驟包括：

找出那些沿著每一軌跡在一圍繞諸時間點之解析範圍內的格子；

對在解析範圍內被找出的每一格子，用對應於該軌跡之音素序列的識別體更新一為該格子維護的表列；

從各格子表列中移除多重識別體；及

決定對應於至少一個具有大於其表列上識別體平均個數識別體之格子的誤差容許範圍。

13. 根據申請專利範圍第12項之方法，其中該找出那些在一解析範圍內之格子的步驟包括處理沿著諸軌跡之諸時間點及更新關聯於在對應解析範圍內的諸格子之表列。

14. 根據申請專利範圍第12項之方法，其中該解析範圍與該誤差容許範圍的大小相同。

15. 根據申請專利範圍第1項之方法，其中該等許多個音標序列中的至少兩個音標序列具有對應於一特別音素序列的部份，該方法尚包括下列步驟：

根據對應軌跡接近誤差容許範圍的程度決定該等音標序列之各部份的值，其中該特別音素序列之聲音元件係根據該等被決定的值從該等音標序列之諸對應部份中的一個形成。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 六、申請專利範圍

16. 根據申請專利範圍第15項之方法，其中該決定諸值的步驟尚根據該對應音標序列的品質衡量值。
17. 根據申請專利範圍第16項之方法，其中該品質衡量值係由一軌跡接近一對應於一不同的邊界性音素之音標序列的誤差容許範圍之程度決定。

18. 一種產生合成語音的裝置，該裝置包括一聲音元件資料庫，該資料庫包含用來連貫以產生合成語音的聲音元件，該聲音元件資料庫藉下列諸步驟建立：

對於對應於包含在發生於一語音信號的一期間內之許多個音標序列中的特定音節的至少一個音素而言，

根據對應於和一誤差容許範圍交叉的不同音素序列之音標序列的軌跡集中度，決定一表示空間內一誤差容許範圍之相對位置，其中每一軌跡代表包含該特定音節之相關音標序列的至少一部份之聲音特性；及

根據諸時間點與誤差容許範圍接近的程度，藉著找出沿著對應軌跡上各相關時間點處之音標序列內的切斷點，由音標序列形成聲音元件。

19. 根據申請專利範圍第18項之裝置，其中該表示空間為一包括許多個連續N維格子的N維空間，且其中該決定誤差容許範圍之步驟包括：

找出那些沿著每一軌跡在一圍繞諸時間點之解析範圍內的格子；

對在解析範圍內被找出的每一格子，若一對應於該軌跡之音素序列的識別體未出現於該格子的表列內，則用

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 六、申請專利範圍

該識別體更新一為該格子維護的表列；及

決定對應於至少一個具有大於其表列上識別體平均數個識別體之格子的誤差容許範圍。

20. 根據申請專利範圍第19項之裝置，其中該找出那些在解析範圍內之格子的步驟包括處理沿著諸軌跡之諸時間點及更新關聯於在對應解析範圍內的諸格子之表列。

21. 根據申請專利範圍第18項之裝置，其中該表示空間為一包括許多個連續N維格子的N維空間，且其中該決定誤差容許範圍之步驟包括：

找出那些沿著每一軌跡在一圍繞諸時間點之解析範圍內的格子；

對在解析範圍內被找出的每一格子，用對應於該軌跡之音素序列的識別體更新一為該格子維護的表列；

從各格子表列中移除多重識別體；及

決定對應於至少一個具有大於其表列上識別體平均個數識別體之格子的誤差容許範圍。

22. 根據申請專利範圍第21項之裝置，其中該找出那些在解析範圍內之格子的步驟包括處理沿著諸軌跡之諸時間點及更新關聯於在對應解析範圍內的諸格子之表列。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

86年3月18日修正補充

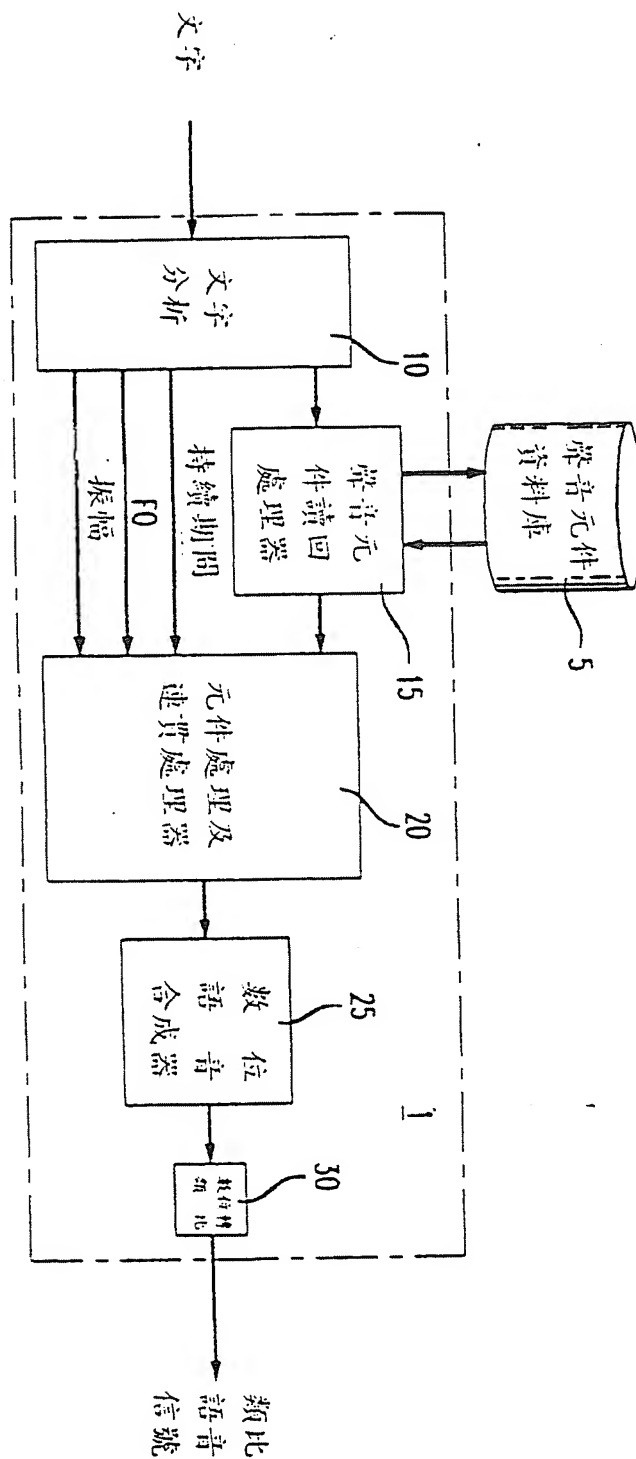


圖 1

図 2A

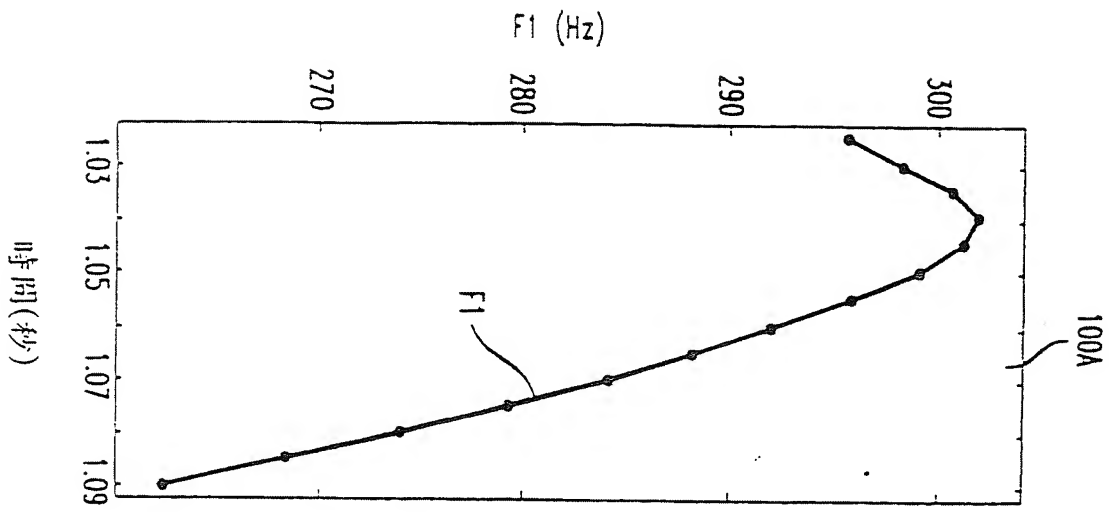


図 2B

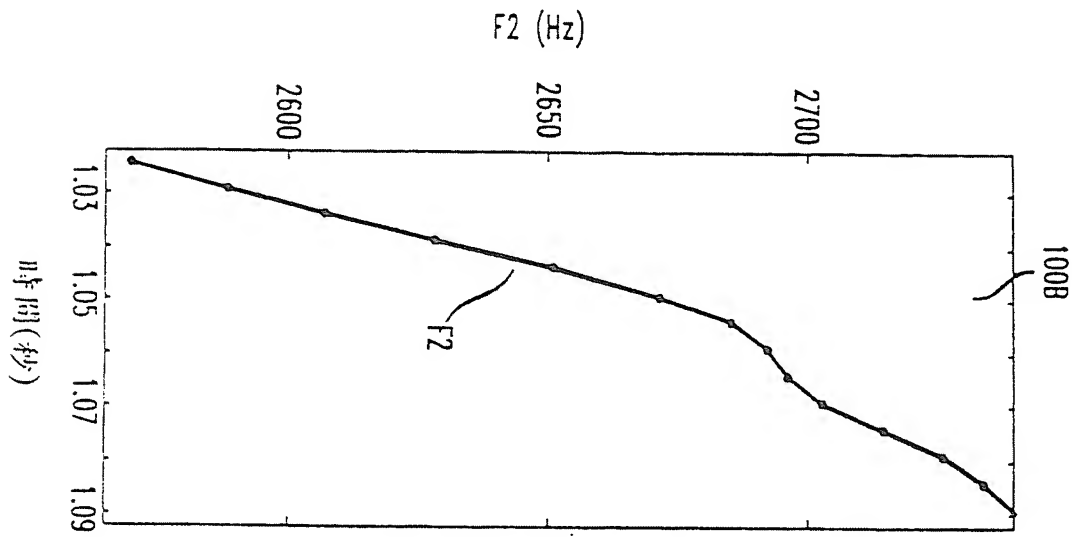


図 2C

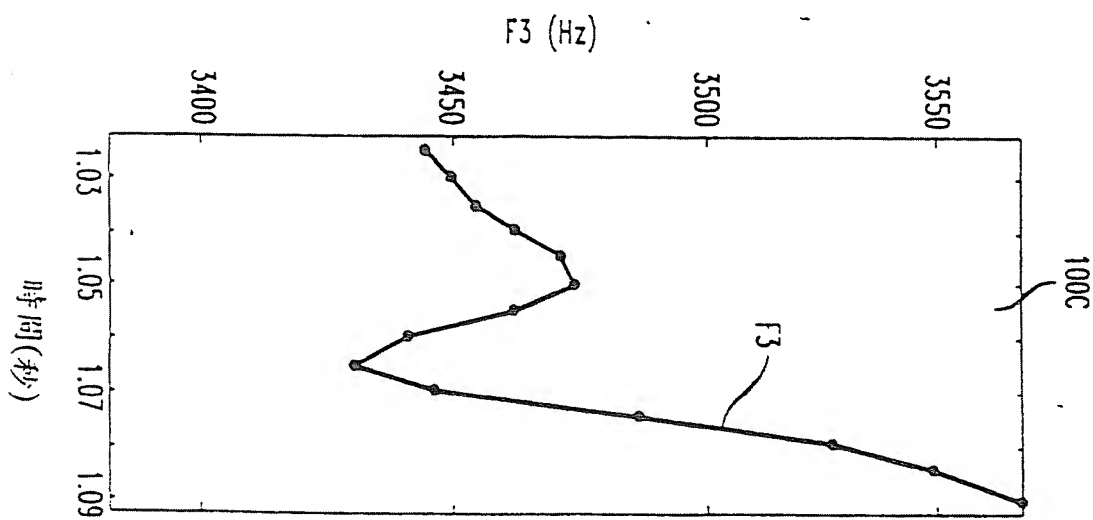




圖 3

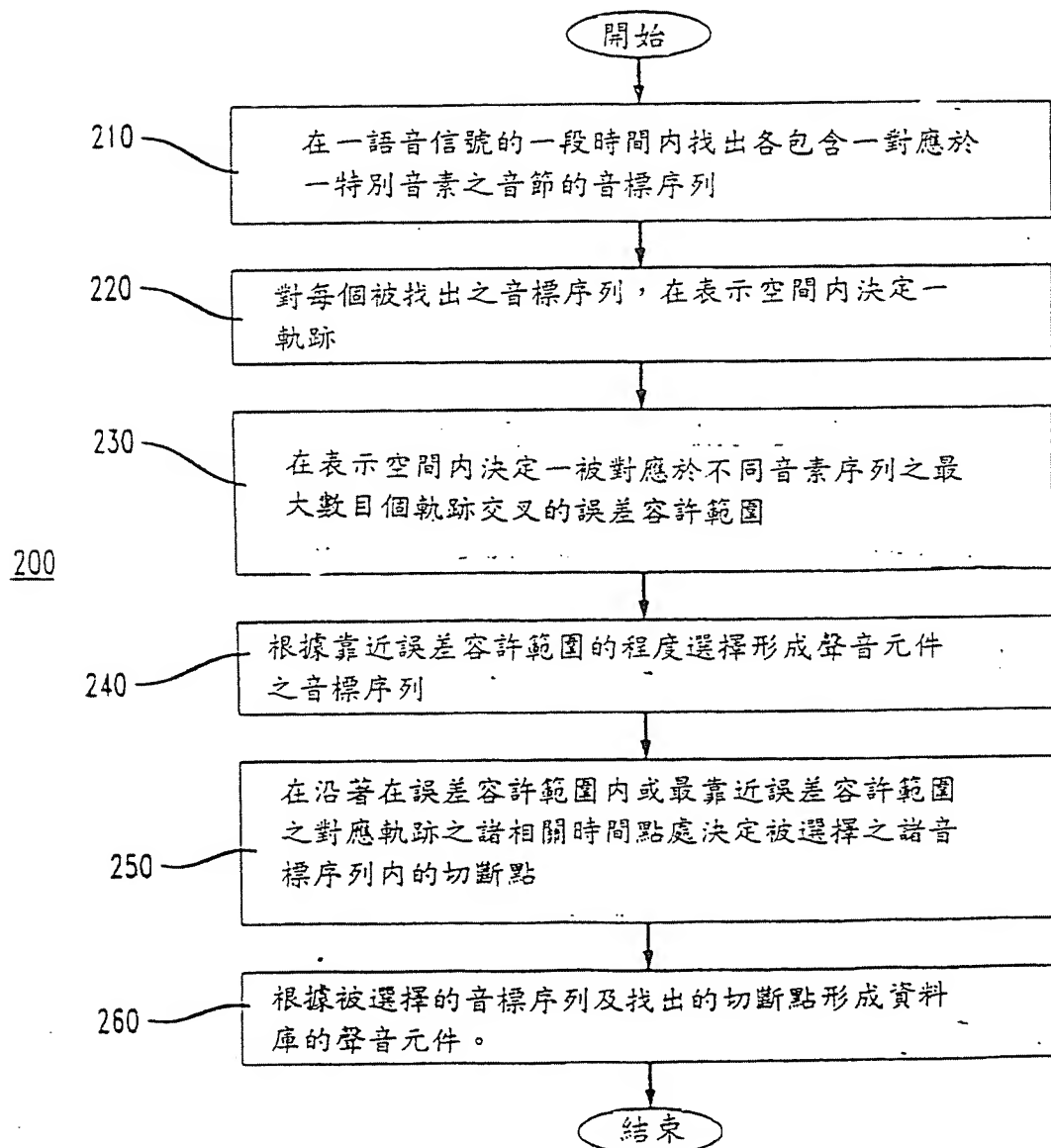


圖 4

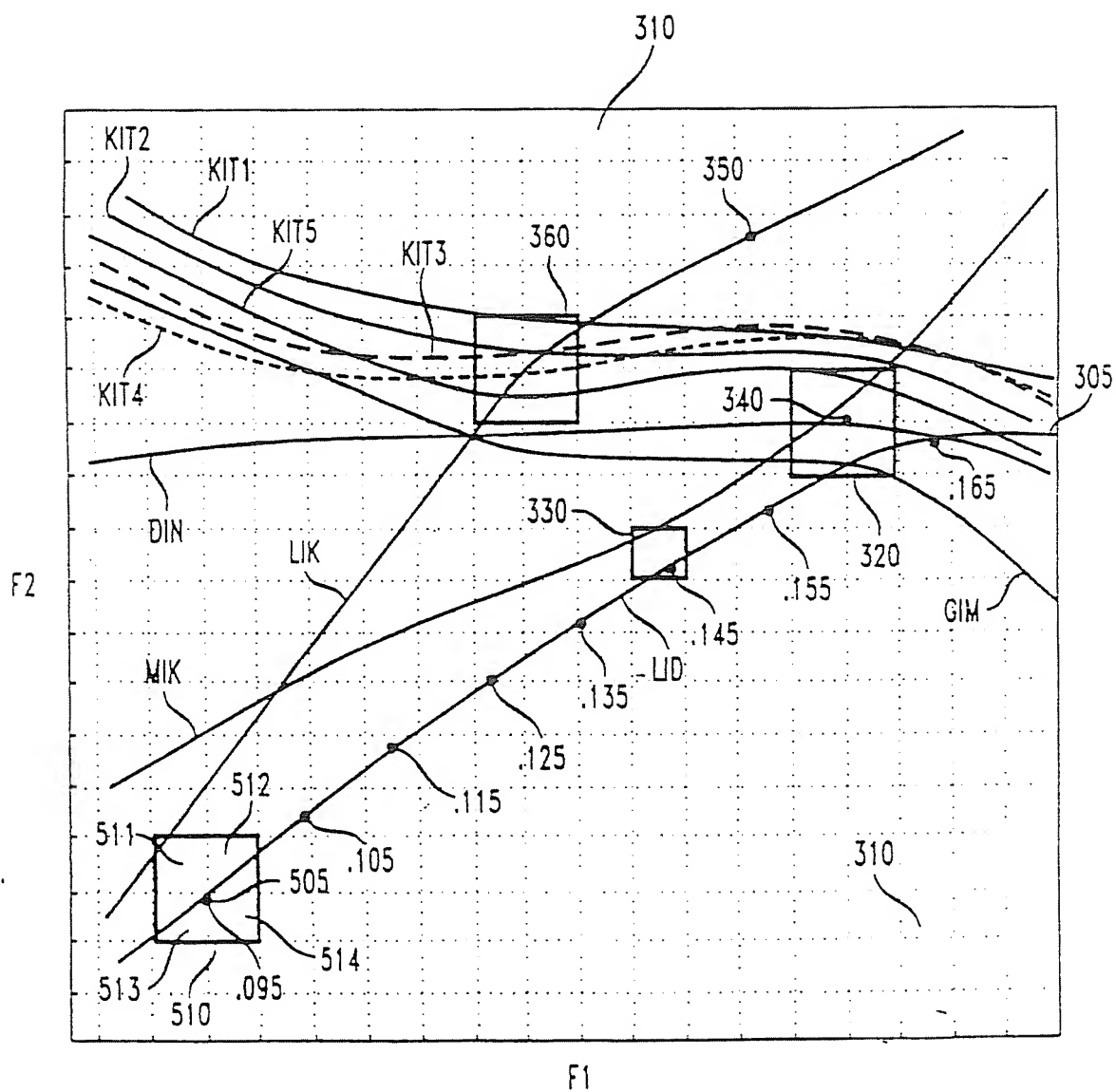


圖 5

